

## Оптимальное распараллеливание для задач моделирования многофазных сред

ПЕКУНОВ В.В., канд. техн. наук

Рассматриваются проблемы оптимального распараллеливания при численном моделировании многофазных сред. Предлагается схема выбора оптимальной комбинации методов распараллеливания. Предложены стратегии сокращения и оптимизации обменов данными при распараллеливании по пространству. Часть обменов исключена за счет периодической экстраполяции с подстройкой весовых коэффициентов в целевой функции метода наименьших квадратов. Для оптимизации оставшихся обменов решается специальная задача нелинейной целочисленной оптимизации.

*Ключевые слова:* моделирование многофазных сред, распараллеливание, метод наименьших квадратов.

## OPTIMAL PARALLELING FOR PROBLEMS OF MULTIPHASE ENVIRONMENT MODELING

PEKUNOV V.V., Ph.D.

The article deals with the problem of optimal paralleling during computational modeling of multiphase environment. It contains the scheme of choice of paralleling method optimal combination. It suggests the strategies of data transmission reduction and optimization while paralleling expansion. Some transmissions are eliminated due to periodic extrapolation with weight ratio tweaking in least-squares method objective function. To optimize the transmissions left a special problem of nonlinear integer optimization.

*Key words:* multiphase environment modeling, paralleling, least-squares method.

Модели многофазных сред обычно включают системы параболических и эллиптических дифференциальных уравнений, записанных для двух или трех пространственных координат. Типичным примером является трехфазная модель образования и распространения загрязнений в воздушной среде (фрагменты модели приведены в работах [1, 2]).

Расчет по таким моделям, особенно в трехмерном случае, весьма трудоемок, что приводит к необходимости применения многопроцессорных вычислительных систем. При этом чрезвычайно важна проблема оптимального распараллеливания расчета в соответствии со структурами задачи и вычислителя. Нами предлагается схема выбора оптимальной комбинации методов распараллеливания (РПП – по пространству и РПФ – по функциям) для вычислительных систем с неоднородной средой передачи данных (например, для получивших широкое распространение кластерных систем с многопроцессорными/многоядерными блоками), являющаяся обобщением и усовершенствованием методик, предложенных нами ранее [3]. Также предлагаем стратегии сокращения и оптимизации обменов данными при распараллеливании по пространству.

**Схема выбора оптимальной комбинации методов распараллеливания.** Будем исходить из принципа минимизации общих временных затрат на вычисления и пересылки данных с учетом неоднородности среды передачи в многопроцессорной системе. Типичным примером такой системы является кластер, установленный в Ивановском государственном энергетическом университете, включающий 32 блока (в каждом по четыре процессора, взаимодействующих через общую память), обменивающихся данными через высокопроизводительную разделяемую коммуникационную среду Gigabit Ethernet.

Предлагаем модель целочисленной нелинейной оптимизации с ограничениями, позволяющую осуществить выбор наилучшего метода (РПФ, РПП, комбинированный) для систем с неоднородной средой передачи данных.

Пусть  $j = \overline{1, N_{\Pi}}$ ,  $k = \overline{1, N_{yp}}$ , где  $N_{\Pi}$  — число процессоров,  $N_{yp}$  — число уравнений. Введем целочисленные переменные:

- $X_{kj} \in \{0; 1\}$ , выражающие факт отображения всего поля  $k$ -й переменной на  $j$ -й процессор (РПФ);
- $F_k \in \{0; 1\}$ , определяющие РПП для поля  $k$ -й переменной;
- $G_j \in \{0; 1\}$ , определяющие участие  $j$ -го процессора в обработке переменных с применением РПП.

Целевая функция имеет вид

$$\max_{r \in \overline{1, N_{yp}}} \left[ \frac{V}{\tau_r} \left( 1 - F_r + F_r \sum_{i=1}^{N_{\Pi}} G_i \right) + \frac{F_r V^+}{\tau_r} \right] + \sum_{r=1}^{N_{yp}} \sum_{f=1}^{N_{yp}} \left[ T_{rf}^K(X, F, G) + T_{rf}^{\Pi}(F, G) \right] \rightarrow \min;$$

$$T_{rf}^K(X, F, G) = A_{rf} \left\{ V \sum_{i=1}^{N_{\Pi}} \sum_{j=1}^{N_{\Pi}} Q_{ij} \left[ X_{ri} X_{fj} + \frac{(X_{ri} F_r G_j + X_{fi} F_r G_i)}{\sum_{i=1}^{N_{\Pi}} G_i} \right] \right\};$$

$$T_{rf}^{\Pi}(F, G) = F_r F_f V_r^S \sum_{m=2}^{N_{\Pi}} G_m Q_{mm-1},$$

где  $V$  — объем поля одной переменной, байт;  $\tau_r$  — скорость обработки поля  $r$ -й переменной, байт/с;  $V^+$  — объем дополнительных данных, обрабатываемых процессором при использовании РПП;  $A_{rf}$  — матрица отношения зависимости  $f$ -й переменной от  $r$ -й переменной;  $Q_{ij}$  — матрица значений времени пересылки одного байта данных из  $i$ -го в  $j$ -й процессор (латентностью пренебрегаем, поскольку она существенно меньше времени передачи), с/байт;  $V_r^S$  — объем данных поля  $r$ -й переменной, пересылаемых между двумя соседними процессорами (в одном направлении) при использовании РПП.

Первый член целевой функции определяет общее время параллельного расчета полей пере-

менных. Второй член выражает суммарное время передачи данных. Функция  $T_{гf}^{\Pi}(F, G)$  определяет время передачи для случая, когда обработка г-й и f-й переменных распараллелена по пространству. Время передачи для прочих случаев (РПФ и РПФ, РПФ и РПП) вычисляется функцией  $T_{гf}^K(X, F, G)$ .

Введем следующие *ограничения*:

- $\forall k: \sum_{j=1}^{N_{\Pi}} X_{kj} = 1 - F_k$ , отражающее тот факт,

что любая переменная обрабатывается с использованием только одного метода распараллеливания;

- $\forall j: \sum_{k=1}^{N_{\text{yp}}} X_{kj} \leq 1 - G_j$ , определяющее, что ка-

ждый процессор может участвовать в параллельной обработке с применением лишь одного из методов;

- $G_1 = \sqrt[N_{\text{yp}}]{F_k}$ ;  $G_2 = \sqrt[N_{\text{yp}}]{F_k}$ , говорящие о том,

что если хотя бы одна переменная нуждается в обработке с использованием РПП, то под такую обработку должны быть выделены по меньшей мере два процессора;

- $\forall j > 2: G_{j-1} - G_j \geq 0$ , задающее непрерывность нумерации процессоров, отведенных под обработку с применением РПП.

#### Проблемы оптимизации обменов данными.

При распараллеливании по пространству расчетная область делится на блоки, каждый из которых обрабатывается одним процессором. Для обеспечения непрерывности первой и второй производных блоки перекрываются на несколько слоев узлов. В ходе расчета процессоры на каждой итерации обмениваются данными на стыках блоков для получения актуальных значений переменных [3]. Часто это единственный обмен данными на итерации, который нельзя сделать полностью асинхронным, что не позволяет достичь наивысшей эффективности распараллеливания. Есть два основных пути решения данной проблемы: а) сократить часть обменов за счет периодической экстраполяции во времени значений переменных на стыках; б) построить оптимальную схему обменов, минимизирующую временные затраты.

**Сокращение количества обменов данными.** Данная задача решается динамически, при исполнении программы. Здесь необходим поиск компромисса между точностью и трудоемкостью вычислений. Воспользуемся экстраполяцией полиномами третьей степени. Особую проблему для прогнозирования представляет тот факт, что на первых сотнях итераций в расчетной области идут быстрые переходные процессы. При этом наибольшей точностью обладает прогноз по небольшому количеству предыдущих точек. С увеличением количества итераций число предыдущих точек, необходимых для достаточно точной экстраполяции, обычно возрастает.

Применим метод наименьших квадратов. Проблема учета различного количества предыдущих точек решается с помощью весов  $q_i(t)$ :

$$\sum_{i=1}^N q_i(t) \delta_i^2 \rightarrow \min,$$

где N – количество предыдущих точек;  $\delta_i$  – невязка для i-й точки.

Задача решается аналитически, коэффициенты полиномов находятся индивидуально для каждого узла расчетной сетки путем решения соответствующей системы линейных уравнений методом LU-разложения.

Для вычисления весов приемлемые результаты дает функция-сигмоида, у которой величина разности между максимальным и минимальным значениями уменьшается с течением времени:

$$q_i(t) = \frac{1}{\exp \left[ A \left( 0,5N / \exp \left( \frac{BNt}{t_{\max}} \right) - i \right) \right] + 1},$$

где A, B – эмпирические коэффициенты;  $t_{\max}$  – максимальное время расчета.

Однако наилучшие результаты демонстрируются при периодической подстройке весовых коэффициентов. Для случайного набора из S узлов расчетной области периодически решается (относительно  $q_i(t)$ ) задача

$$\sum_{j=1}^S \sum_{k=1}^M \delta_{jk}^2 \rightarrow \min,$$

где M – глубина предсказания;  $\delta_{jk}$  – невязка в j-м узле для k-й точки.

Как показывает опыт расчетов, подстройка снижает погрешность экстраполяции в 1,2–1,3 раза на наиболее критичном участке – на первых сотнях итераций. Далее погрешность экстраполяции стабильно падает, подстройка дает снижение погрешности не более чем в 1,05 раза.

**Построение оптимальной схемы обменов.** Данная задача решается на этапе составления программы. Совместим во времени вычисления и обмены данными на стыках. Вычислительный процесс представим как серию процессов обработки сильно связанных групп переменных модели. Редуцируем оргграф взаимозависимостей переменных модели, описанный матрицей смежности A, последовательно объединяя узлы, образующие циклы. Получим сеть взаимозависимостей K групп переменных, по которой построим матрицу P[K×K] отношений строгого следования между группами:

$$P_{ij} = \begin{cases} 1, & \text{если } (i\text{-я группа}) \succ (j\text{-я группа}), \\ 0, & \text{иначе.} \end{cases}$$

Совмещение вычислений и обменов ограничивается лишь наличием зависимостей между группами. Обработку некоторой j-й группы можно совместить во времени с асинхронной передачей данных, необходимых для обработки i-й группы, если  $P_{ij} = 1$ .

Введем  $T_S^j(N)$  – таблично заданную нелинейную функцию времени передачи сообщения (на j-й стадии) из N элементов. Необходимо учесть, что если время обработки ( $t_p^j$ ) текущей группы меньше времени асинхронного обмена ( $t_s^j = T_S^j(N^j)$ ) данными для следующей группы, то время выполнения стадии равно  $t_s^j$ . Если таких стадий несколько, то

общее время их выполнения равно сумме величин  $t_S^j$ . Учитывая, что обычно [3]

соответствующие группы целесообразно объединить в блок групп, что позволит сократить время выполнения за счет уменьшения времени обменов. Таким образом, задачу оптимизации обменов логично решать относительно блоков групп.

Следует найти такую непротиворечивую последовательность обработки блоков групп и пересылок данных, которая обеспечивает минимальные затраты времени. Дадим формальную постановку соответствующей задачи *нелинейной целочисленной оптимизации*. Необходимо найти значения целочисленных переменных  $X_{ij} \in \{0, 1\}$ , выражающих факт вхождения  $i$ -й группы в блок групп, передаваемый  $j$ -м по счету и обрабатываемый  $(j+1)$ -м по счету, при которых достигается минимум временных затрат:

$$\sum_{j=1}^M T_S^j(N_j) + \sum_{j=1}^{M-1} \frac{|kW_j - T_S^{j+1}(N_{j+1})| + (kW_j - T_S^{j+1}(N_{j+1}))}{2k} +$$

$$+ W_M \rightarrow \min;$$

$$N_j = \delta \sum_{i=1}^K X_{ij} V_i;$$

$$W_j = \sum_{i=1}^K X_{ij} \frac{V_i}{\tau_i};$$

$$j = \overline{1, M},$$

где  $M$  – максимальное количество блоков групп;  $k$  – эмпирический коэффициент замедления вычислений при асинхронном обмене;  $\delta$  – доля объема данных блока поля, подлежащая передаче;  $\tau_i$  – эмпирическая средняя скорость обработки поля одной переменной в  $i$ -й группе;  $V_i$  – объем данных в  $i$ -й группе.

В целевой функции учтены: а) возможность как полного, так и частичного совмещения вычислений и обменов; б) нетривиальный факт, что при совмещении вычислений и пересылок скорость обработки данных может уменьшаться с некоторым коэффициентом  $k$  (по техническим причинам, например, при появлении конфликтов на шине данных вычислительного узла).

Зададим отсутствие простоев между обработкой блоков групп, вхождение любой группы лишь в один блок и непротиворечивость порядка обработки и взаимозависимостей групп:

$$\forall j > 1: \bigvee_{i=1}^K X_{ij-1} - \bigvee_{i=1}^K X_{ij} \geq 0;$$

$$\forall i: \sum_{j=1}^M X_{ij} = 1;$$

Пекунов Владимир Викторович,

ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина», кандидат технических наук, доцент кафедры высокопроизводительных вычислительных систем, телефон (4932) 26-98-29.

$$T_S^j(\sum N^j) \leq \sum T_S^j(N^j),$$

$$\forall i, j: K(1 - X_{ij}) \geq \sum_{r=1}^{j-1} \left( \sum_{s=1}^K X_{sr} P_{si} \right).$$

Таким образом, предложенные схемы позволяют добиться оптимального распараллеливания и обладают достаточной общностью для решения задач с моделями, включающими параболические и эллиптические уравнения в частных производных:

а) схема выбора оптимальной комбинации методов распараллеливания для многопроцессорных систем с неоднородной средней передачи данных;

б) стратегия сокращения количества обменов данными на стыках блоков расчетной области при распараллеливании по пространству;

в) нелинейная оптимизационная модель, позволяющая определить оптимальную схему обменов данными на стыках.

На базе предложенных схем произведен выбор оптимальной комбинации методов распараллеливания для МВС-1000/16. Показано, что для нашей задачи наиболее оптимальным является применение РПП для всех переменных. Для сокращения и оптимизации обменов использовались предложенные стратегии. Были проведены численные эксперименты по моделированию распространения различных видов загрязнений. При  $N = 5$  и  $M = 4$  в системе МВС-1000/16 удалось повысить эффективность распараллеливания с 80% до 91% на девяти процессорах, с 85,5% до 93% на пяти процессорах, с 86% до 90% на двух процессорах.

Работа была выполнена при финансовой поддержке Минобразования и науки (грант РНП.2.2.1.1.7280).

#### Список литературы

1. Пекунов В.В., Ясинский Ф.Н. Математическая модель микроклимата в производственных помещениях с повышенной влажностью // Изв. вузов. Технология текстильной промышленности. – 2006. – № 2. – С. 128–133.
2. Пекунов В.В. Учет фактора излучения при моделировании процесса образования и распространения загрязнений в воздушной среде // Вестник ИГЭУ. – 2006. – Вып. 3. – С. 76–79.
3. Пекунов В.В., Ясинский Ф.Н. Параллельное решение задачи численного моделирования распространения загрязнений в воздушном бассейне большого города и в окрестности предприятия: Препринт № 36 / Институт прикладной математики им. М.В. Келдыша РАН. – М., 2003.